(12) # EUROPEAN PATENT APPLICATION

(72) Inventor : Nakabayashi, Jiro
1-15-1-606, Tajima
Urawa-shi, Saitama (JP)

(74) Representative : Brown, Kenneth Richard et al
R.G.C. Jenkins & Co.
26 Caxton Street
London SW1H 0RJ (GB)

(54) Audiovisual communication system using variable length packets.

(57)   The purpose of the invention is to provide a system capable of easily reproducing an image from interframe predictive signals conventionally used for, e.g., visual telephone communications. Video and audio signals are multiplexed and assembled into packets with lengths proportional to the bit rates of the multiplexed data by a packet separating-multiplexing means (59) and completed with information on the multiplexing conditions, i.e., packet structure, then stored and when the stored packets are reproduced, the packet separating-multiplexing means judges the structure of the packets according to the information of the multiplexing conditions and disassembles the packets into individual data and decodes the video and sound signals.

FIG.6

EP 0 601 766 A1
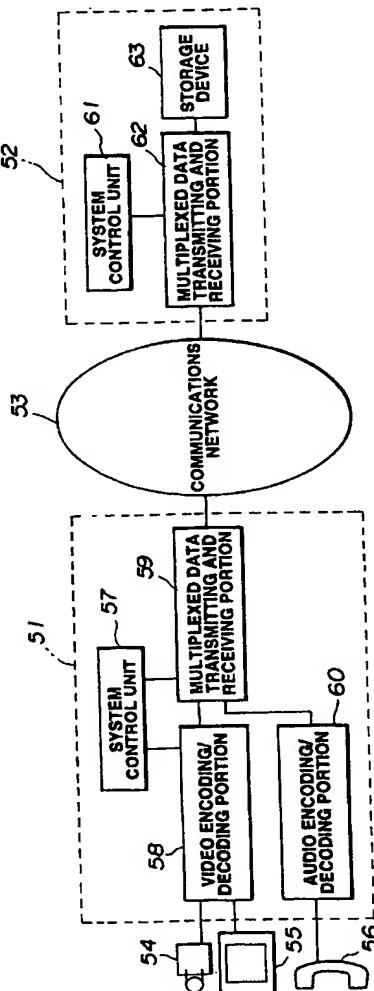
BACKGROUND OF THE INVENTION

The present invention relates to a video communications system and more particularly to a video communications system capable of compressing information that represents an image, particularly, a dynamic image i.e. a moving picture, of multiplexing and storing the information and audio data and finally of regenerating and decoding the stored information.

Recently, in addition to conventional communications services, i.e. telephone and facsimile, audio storage and delivery services have been available over certain networks to allow audio data to be stored temporarily until a particular destination terminal was available.

The progress in digital communications networks, i.e. ISDN (Integrated Services Digital Networks) enables the terminal users to receive video communications services by means of visual telephone terminals. With such a background, there is an increasing demand for video storage and delivery services which are put into practice with certain local applications.

In a conventional video communications system, an input video signal through the camera of the transmitting terminal enters into the encoding portion whereby it is encoded with stored amounts of information by an interframe predictive coding method. On the other hand, an audio signal input through the handset is encoded by the audio encoding portion. Both encoded video and audio signals enter into the multiplexing portion whereby they are combined with each other and completed with a video control signal designated by the terminal system control unit after which they are sent to the communications network. The multiplexed data is transmitted over the network to the remote terminal wherein the data is separated into individual signals i.e., a video signal, an audio signal and a video control signal by the action of multiplexing portion. The video control signal, the video coded signal and the audio coded signal are transferred to the system control unit, the video decoding portion and the audio decoding portion, respectively. In the video decoding portion, the video data encoded by the interframe predictive coding method is decoded and then is put out to create an image on the screen of the monitor. In the audio decoding portion, coded audio data is decoded and put out to the handset. The video control signal is treated by the system control unit. The above-mentioned actions are conducted through a two-way transmission line to produce bidirectional visual telephone communications.

The interframe predictive coding method is such that video data of a predicted picture is encoded in a compressed form to include only differential data between a standard updated picture and the succeeding picture, utilizing a very strong correlation between the frames of the moving picture data. This method enables us to considerably reduce the number of video data needed to be transmitted since the predicted picture includes a very small amount of data in comparison with the updated picture. The header information is attributive information attached, when coding, to the beginning of the corresponding frame of video data. The updated picture is one frame of video data encoded therein in the same way as a still picture, and after coding, it has the number of the coded data equivalent to that of the still picture. The interframe predicted picture is defined by determining the a difference between the updated video data and the video data of a subsequent frame and it is represented as a frame which includes only differential data.

In a conventional video storing and reproducing system utilizing still pictures, the host and the terminal are connected to a communications network (such as ISDN) and operate to gain access to each other at the host's or the terminal user's request. When a circuit therebetween is established, a guiding picture from the host is displayed on the screen of the monitor. Referring to the guiding display, the terminal user selects a menu for storing or reproducing video signals by one frame representing a still image with a audio signal. The video signal to be stored in the memory is input through the camera and an audio signal to be stored is input through the handset. The input video and audio signals are encoded respectively by the video encoding portion and the audio encoding portion and then they are transmitted over the communications network to the host which in turn combines and stores the received encoded video and audio signals in the storage device .

A set of data is described in such a way that a packet of one-frame video data is combined with the audio data having an identifying flag and the character data also having an identifying flag. The number of bits in a packet of video data depends upon the types of video signals to be encoded but it may not largely vary since coding is achieved within one frame. Therefore, the audio data has no considerable variation in its length. Sometimes an image may contain character data, and sometimes it may contain none. However, the amount of character data per scene is much smaller than that of the video data and, therefore, it is negligible.

The prior art for which the present invention was made is disclosed in the technical report titled "CODEC for video delivery services and video Conference services" of IEIEC (Institute of Electronics, Information and Communication Engineers) No. IE93-41, 1993-07. The prior art dealt with the same contents and the same users' system of communication services as the present invention but it proposed that each terminal periodically perform in-frame coding to make the data suitable for storage and reproduction.

On the contrary, the present invention provides

that universal visual telephone terminals be used and that the communications center make the data adaptable for storage and reproduction.

On the other hand, the present invention relates to a data multiplexing system which enables all stored data, including video signals, to be reproduced in conformity with the communications conditions that will assure broad utilization of storage and forward communications services by means of terminals such as universal visual telephone equipment.

As mentioned above, the conventional video delivery services must have center facilities to generate information, change the stored data, convert the media and the stored data by half. On the other hand, the data multiplexing method currently used in video communications, i.e. visual telephone communications, combines the data of each half by a bit. If the multiplexed data is stored as is, the center processor can hardly generate any information or change it.

In the case of realizing video delivery services including dynamic images, the video transmitting format currently applied, for example in visual telephone services, utilizes the interframe prediction technique for compressibly encoding the video data and therefore it may have a certain amount of information per frame which considerably differs from the format utilizing the interframe predictive coding. Consequently, when packets for storing video and audio data according to video signals per frame are formed, the packets may differ from one another in the size and the amount of information to be stored. This may cause the discontinuation of the data's flow when reproducing the stored video data in a certain sequence. Since the conditions of the communications between the terminals may be changed at any time, it is impossible to use a packet of a fixed structure.

The conventional coding method may be applied by increasing the number of frames of still images without using an interframe predictive coding to simulate dynamic images. In this case, the quantity of data encoded without interframe prediction is ten times larger than that of the data encoded by the interframe predictive coding. An attempt to transmit data necessary for regenerating an image to be observed as it moves, leads to an increase in the cost of the transmission. This may also require an increase in the capacity of a transmission line.

SUMMARY OF THE INVENTION

It is an object of the present invention to provide a video communication system which is made to provide a video communication system for use in dynamic video (i.e. moving pictures) store and delivery services and which is capable of easily storing and- reproducing conventionally used video signals, i.e. visual telephone signals encoded by interframe predictive coding thereby eliminating the possibility of an interruption in the flow of data during in the process of image reproduction.

It is another object of the present invention to provide a video communication system, when the signals of a frame are separated and stored, separated video data, audio data and character data are assembled into respective packets which correspond to the respective bit rates and are completed each with a header of information identifying the packet's structure and the multiplexing conditions and is then stored. When reproducing the stored packets, the structure of each packet is judged from the header information and the stored data is reassembled from the packets and reproduced. Therefore, the data separating-multiplexing system, according to the present invention can achieve easy storage/reproduction of interframe predictive video signals such as those adopted in visual telephone services. This system can also realize the continuous data reproduction of moving pictures, even at conventional visual telephone terminals.

The object of the present invention is to provide a video communications system wherein video and audio signals multiplexed by the first separating-multiplexing means are assembled into respective packets, each respective packet being of a length proportional to the bit rate of data for the respective signal type, by a packet separating-multiplexing means, completed with information on the packet's structure and then stored in the form of a series of packets. When the stored packets are reproduced, the packet separating-multiplexing means analyzes the structure of each packet according to the information of the multiplexing conditions and disassembles the packets in to individual data and decodes the video and audio signals.

The application of the above-mentioned data multiplexing and separating system enables the communications center to prepare and/or change the data and to store data, encoded by the conventional dynamic video coding method and to complete the video-delivery services without an interruption of data in its reproduction at any of the conventional visual telephone terminals.

To realize the above-mentioned objectives, the present invention was made:

To provide a video communications system comprising a first CODEC (converting means) to encode or decode input video data; a second CODEC (converting means) to encode or decode inputted audio data; a first separating-multiplexing means to separate and multiplex signals input from the first conversion means and the second conversion means and other data; a first communications means to transmit signals multiplexed by the first separating-multiplexing means; a first control means to control the first converting means, the second converting means, and the first communications means; a second com-

munication means to receive the multiplexed signals transmitted from the first communications means; a second separating-multiplexing means to separate a part of the signals received by the second communications means; a packet separating-multiplexing means to assemble data into packets to be stored; a storage means to store the data packets formed by the packet separating-multiplexing means; a second control means to control the second communications means, the second separating-multiplexing means, and the packet separating-multiplexing means; characterised in that video and audio signals multiplexed by the first separating-multiplexing means are assembled into packets, each respective packet being of a length proportional to the bit rate of data for the respective signal type, by the packet separating-multiplexing means and completed with information on the multiplexing conditions, i.e. the packets are formed; and then thus the packets are stored and when the stored packets are reproduced, the packet separating-multiplexing means identifies the packet's structure according to the information of the multiplexing conditions, disassembles the packets into separate data and further decodes the video and audio signals.

BRIEF DESCRIPTION OF DRAWINGS

Figs. 1A and 1B shows the conventional format of multiplexed data.

Fig. 2 is a construction view of a conventional video communications system.

Fig. 3 shows a data format for use in a conventional video communications system.

Fig. 4 is a construction view of a conventional video storing and reproducing system utilizing still pictures.

Fig. 5 shows a stored data format for a conventional video storing and reproducing system utilizing still pictures.

Fig. 6 is a construction view for explaining a video communications system embodied in the present invention.

Fig. 7 is a construction view of a separating and multiplexing portion of the multiplexed data transmitting and receiving portion shown in Fig. 6.

Fig. 8 shows the format of data multiplexed by a data separating and multiplexing system according to the present invention.

PREFERRED EMBODIMENT OF THE INVENTION

Figs. 1A and 1B show a conventional data multiplexing format for use in multiplexing data for visual telephone communications and the like. This system is adopted as an international standard for data multiplexing by which the interconnection of visual telephones or the like is assured. In the drawings, numer-

als 1, 2 and 3 designate areas allocated to audio data, character data and video data, respectively. Numeral 4 denotes bits for frame synchronization, numeral 5 denotes bits for transmitting capacity information and commands between terminals; and numeral 6 denotes a sequence for transmitting data to a communication channel. Data transmission begins with the first bit of the first acted. It is always possible to change the multiplexing condition of the information by using the bits 5 for transmitting the capacity and commands.

Fig. 2 shows a construction view of a conventional video communications system for use with visual telephones which includes a transmitting terminal 11; a remote terminal 12, a communications network 13, such as ISDN, cameras 14a and 14b to input video signals, monitors 15a and 15b to put out video signals, handsets 16a and 16b to input/output audio signals, terminal system control units 17a and 17b, video encoding/decoding portions 18a and 18b, multiplexing portions 19a and 19b for combining (multiplexing) encoded video data, audio data and control information and the transmitting of the combined coded data over the network and audio encoding/decoding portions 20a and 20b.

An input video signal through the camera 14a of the transmitting terminal 11 enters into the encoding portion 18a whereby it is encoded with stored amounts of information by an interframe predictive coding method. On the other hand, an audio signal input through the handset 16a is encoded by the audio encoding portion 20a. Both encoded video and audio signals enter into the multiplexing portion 19a whereby they are combined with each other and completed with a video control signal designated by the terminal system control unit 17a after which they are sent to the communications network 13. The multiplexed data is transmitted over the network to the remote terminal 12 wherein the data is separated into individual signals i.e., a video signal, an audio signal and a video control signal by the action of multiplexing portion 19b. The video control signal, the video coded signal and the audio coded signal are transferred to the system control unit 17b, the video decoding portion 18b and the audio decoding portion 20b, respectively. In the video decoding portion 18b, the video data encoded by the interframe predictive coding method is decoded and then is put out to create an image on the screen of the monitor 15b. In the audio decoding portion 20b, coded audio data is decoded and put out to the handset 16b. The video control signal is treated by the system control unit 17b. The above-mentioned actions are conducted through a two-way transmission line to produce bidirectional visual telephone communications.

Fig. 3 shows the data format used in a conventional video communications system, as for example, the data format used in the case when video data is

encoded by the interframe predictive coding method. The data format is composed of header information 21a, 21b, 21c, 21d, 21e for identifying the beginning of one frame of video data, to update video data 22 that serves as a standard for interframe predictive coding and video data 23a, 23b, 23c, 23d encoded by the interframe predictive coding method.

Fig. 3 shows the video data encoded by the video decoding portion 18a or 18b according to the interframe predictive coding method. The interframe predictive coding method is such that video data of a predicted picture is encoded in a compressed form to include only differential data between a standard updated picture and the succeeding picture, utilizing a very strong correlation between the frames of the moving picture data. This method enables us to considerably reduce the number of video data needed to be transmitted since the predicted picture includes a very small amount of data in comparison with the updated picture. The header information 21a, 21b, 21c, 21d, 21e is attributive information attached, when coding, to the beginning of the corresponding frame of video data. The updated video data 22 is one frame of video data encoded therein in the same way as a still picture, and after coding, it has the number of the coded data equivalent to that of the still picture. The interframe predicted video data 23a, 23b, 23c, 23d are defined by determining the a difference between the updated video data and the video data of a subsequent frame and it is represented as a frame which includes only differential data.

Fig. 4 is a construction view of a conventional video storing and reproducing system utilizing still pictures. This system is adopted in still video storage and delivery services. The system includes a terminal 31, a communications network 32 (such as ISDN), a communications host 33, video storage device 34 for storing pictorial images, a video input camera 35, a video output monitor 36, a handset 37 for inputting/outputting audio signals, a video encoding/decoding portion 38 for encoding/decoding still picture data and an audio encoding/decoding portion 39 for encoding/decoding audio signals.

The host 33 and the terminal 31 are connected to a communications network 32 (such as ISDN) and operate to gain access to each other at the host's or the terminal user's request. When a circuit therebetween is established, a guiding picture from the host is displayed on the screen of the monitor 36. Referring to the guiding display, the terminal user selects a menu for storing or reproducing video signals by one frame representing a still image with a audio signal. The video signal to be stored in the memory is input through the camera 35 and an audio signal to be stored is input through the handset 37. The input video and audio signals are encoded respectively by the video encoding portion 38 and the audio encoding portion 39 and then they are transmitted over the

communications network to the host which in turn combines and stores the received encoded video and audio signals in the storage device 34.

Fig. 5 shows the data format for storage in a conventional video storing and reproducing system utilizing still images. The system is adopted in a still-video delivery service. The format includes audio data 41a and 41b each having an identifying flag a_1, b_1, a one-frame video data 42a and 42b each having an identifying flag a_2, b_2 and character data 43a and 43b each having an identifying flag a_3, b_3. The data format shown in Fig. 5 is used in the case when the video and audio data are multiplexed and stored of the storage device 34 in the system shown in Fig. 4. A set of data is described in such a way that a packet of one-frame video data 42a (42b) is combined with the audio data 41a (41b) having an identifying flag and the character data 43a (43b) also having an identifying flag. The number of bits in a packet of video data 42a (42b) depends upon the types of video signals to be encoded but it may not largely vary since coding is achieved within one frame. Therefore, the audio data 41a and 41b has no considerable variation in its length. Sometimes an image may contain character data, and sometimes it may contain none. However, the amount of character data per scene is much smaller than that of the video data and, therefore, it is negligible.

Referring now to the accompanying drawings, preferred embodiments of the present invention will be described in detail as follows:

Fig. 6 is a construction view for explaining a video communications system embodied in the present invention. The shown system includes a terminal 51, a host 52, a communications network 53 (such as ISDN), an image inputting camera 54, an outputting image monitor 55, a sound input/output handset 56, a terminal systems control unit 57, a video encoding/decoding portion 58, a multiplexed data transmitting/receiving portion 59 for combining encoded video data, encoded audio data and to control the information and transmitting of the multiplexed data to a line of the communications network, an audio encoding/decoding portion 60, a host system control unit 61, a multiplexed data transmitting/receiving portion 62 for separating the controlled information from multiplexed data and for combining coded data at the host, and a storage device 63 for storing the data.

A dynamic image (moving picture) input through the camera 54 at the terminal 51 is encoded by the video encoding portion 54 according to an interframe predictive coding method. At this time, coded header information is added to the encoded video data according to instructions from the terminal system control unit 57. Audio data input through the handset 56 are encoded by the audio encoding/decoding portion 60. Both coded video and audio data and a control code are multiplexed by the multiplexed data trans-

mitting portion 59 and then transferred to the host 52 through the communications network 53. The host 52 receives the multiplexed data from the terminal 51 at its multiplexed data receiving portion 62 and stores them in the adequate file form in the storage device 63.

The stored data is read out from the storage at the request of the terminal user, or the host, and transferred to the multiplexed data transmitting portion 62 whereby the controlled information of the data is renewed according to the instructions given by the system control unit 61 and them transmitted through the network 53 to the terminal 51. The terminal 51 receives the data having the renewed control code from the host by its multiplexed data receiving portion 59 which separates the data into the control code, the video data and the audio data and transfers the control code to the terminal system control unit 57. The encoded video data separated at the multiplexed data receiving portion 59 is transferred to the video decoding portion 58 wherein the separated video data is decoded according to the instructions of the system control unit 57 and the original moving image is reproduced on the monitor screen 55. The separated audio data is decoded and reproduced by the audio decoding portion 60 and the sound is emitted from the handset 56.

During the storing and reproducing of data, communications may be always realized through a bidirectional channel, allowing any controlled information to be transmitted therethrough irrespective of the existence or absence of video data to be transmitted.

Fig. 7 is a construction view of a data separating/multiplexing portion included in the multiplexed data transmitting/receiving portion (see Fig. 6) according to the present invention. In Fig. 7, numeral 71 designates data multiplexed from audio sources, video sources and character sources and being transmitted through a channel. Numeral 72 indicates a transmission data separating-multiplexing portion for separating the multiplexed data into audio data, numerals 73, 74 and 75 denote, respectively, the audio data, the video data and the character data separated from each other from the data previously encoded and multiplexed at a terminal. Numeral 76 designates a stored data separating-multiplexing portion wherein data separation and multiplexing according to the present invention is carried out and numeral 77 indicates stored data multiplexed by the data separating-multiplexing method according to the present invention.

As shown in Fig. 7, the multiplexed data 71 is transmitted through a bidirectional channel from a terminal to a host and vice versa. The multiplexed data of the format shown in Fig. 1 is inputted in turn bit by bit in, to the transmission data separating-multiplexing portion 72 wherein they are separated into separate audio 73, video 74 and character 75 data.

The separated data is transferred to the storage data separating-multiplexing portion 76 which defines the data multiplexing conditions by using bits 5 for the transimitting of capacity information and which commands in such a way that it may be understood by the data separating-multiplexing portion 72. Then that portion creates multiplexed data 77 and stores it in a storage device 63. The stored data may be transmitted from the host to the terminal in the reverse order of processing.

Fig. 8 shows a multiplexed data format used in a data separating-multiplexing system according to the present invention. In Fig. 8, numerals 81, 82, 83 designate, respectively, an audio data length per packet, a video data length per packet and a character data length per packet. Numeral 84 designates numbers of packets of audio, video and character data and numerals 85, 86 and 87 indicate respectively, an audio data packet, a video data packet and a character data packet.

The data's length 81, 82, 83 for each packet is determined by the calculation from a bit rate of the capacity and command transmission bits 5 and the time of one packet to be formed. The number of packets 84 is equal to the total of the audio, video and character packets to be stored by one storage process. Each kind of data is described in the number of bits designated by the respective data's length 81, 82, 83 and is arranged in a packet 85, 86, 87. Thus a designed packet may have a constant size adapted to the reproducing sequence irrespective of the amount of video data.

When reproducing the stored data, it is possible to judge the bit rate of each kind of stored data from the data's length 81, 82, 83. Consequently, it becomes possible to set the communications' conditions according to the data bit rate by using the capacity and command bits 5, Thereby the data stored under different communications' conditions can be easily reproduced.

Claims

1. A video communications system for compressing information contained in a dynamic image, multiplexing and storing the compressed information together with audio signals and other data, reproducing and decoding the stored information, characterized by; a video communications system comprising: a storage means to assemble video and audio signals into respective packets, each respective packet being of a length proportional to the bit rate of data for the respective signal type, to add a header indicating the multiplexing conditions under which the packets were formed and to store the packets therein; and a judging means to identify the packet's structure

according to the header when reproducing the packets stored in the storage means to enable further disassembly of the packets into separate data, and to decode the video and audio signals.

2. A video communications system for compressing information contained in a dynamic image, multiplexing and storing the compressed information together with audio signals and other data, reproducing and decoding the stored information, characterized by; a video communications system comprising: a first converting means (58) to encode or decode inputted video data; a second converting means (60) to encode or decode inputted audio data; a first separating-multiplexing means to separate and multiplex signals inputted from the first converting means (58) and the second converting means and other data; a first communications means (59) to transmit signals multiplexed by the first separating-multiplexing means; a first control means (57) to control the first converting means (55), the second converting means (60), the first separating-multiplexing means, and the first communications means (59); a second communications means (62) to receive the multiplexed signals transmitted from the first communication means (59); a second separating-multiplexing means to separate a part of the signals received by the second communication means (62); a packet separating-multiplexing means to assemble data into packets to be stored; a storage means (63) to store the data packets formed by the packet separating-multiplexing means; a second control means (61) to control the second communication means (62), the second separating-multiplexing means, and the packet separating-multiplexing means; characterized in that the first separating-multiplexing means assembled the video and audio data into reapective packets, each respective packet being of a length proportional to the bit rate of data for the respective signal type, and heads the packets with information as to the multiplexing conditions under which the packets were formed.

3. A video communication system as defined in Claim 2, characterized in that when the packets stored in the storage means (63) are reproduced, the packets separating-multiplexing means identifies the packet's structure according to the information in the multiplexing conditions, disassembles the packets into separate data and further decodes the video and audio signals.

4. A method of formatting audio visual data for storage and reproduction comprising: arranging the video and audio data in respective packets of lengths proportional to the bit rate of data for each respective data type and heading the packets with a header containing information as to the packet structures.

5. A method as claimed in claim 4 wherein the header contains information as to the length of each packet type.

6. A method as claimed in claim 4 or claim 5, wherein the packets are arranged in an alternating sequence.

7. A method as claimed in any one of claims 4 to 6 wherein the audio visual data includes control data.

FIG.1A
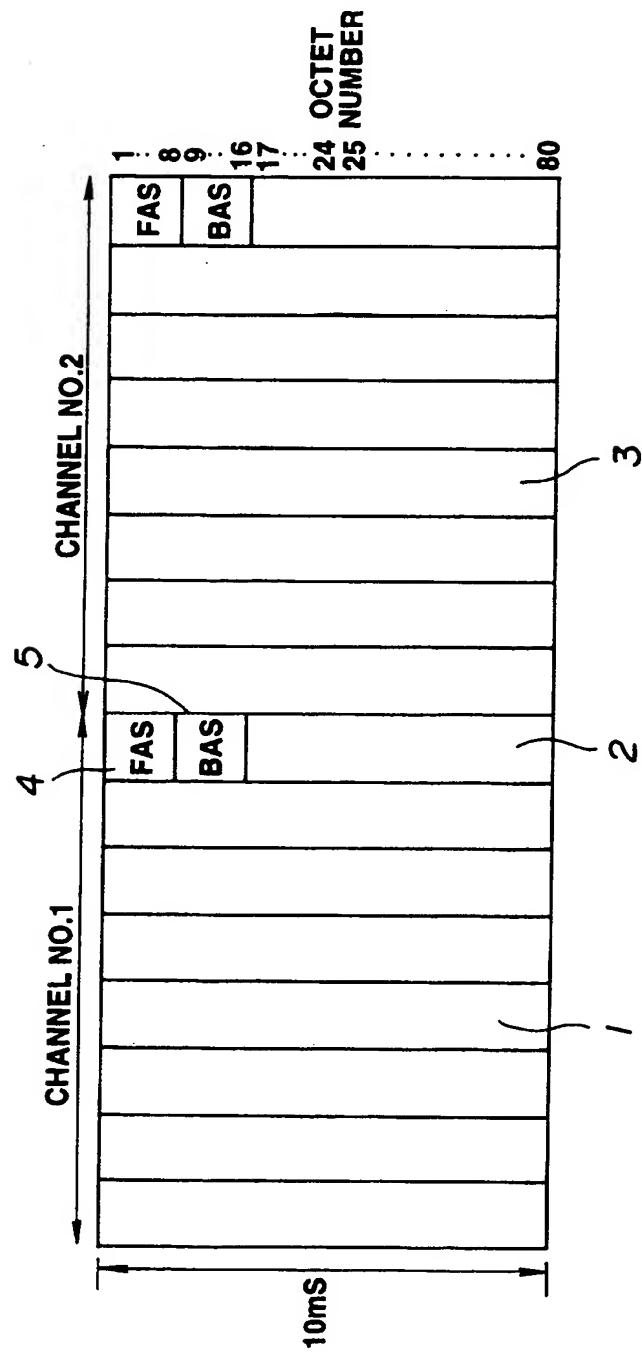
FIG.1B
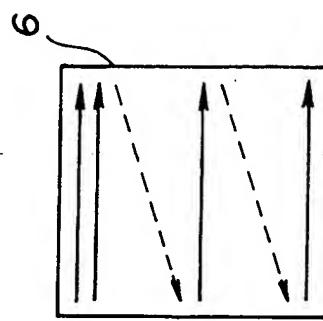
FIG.2

FIG.3
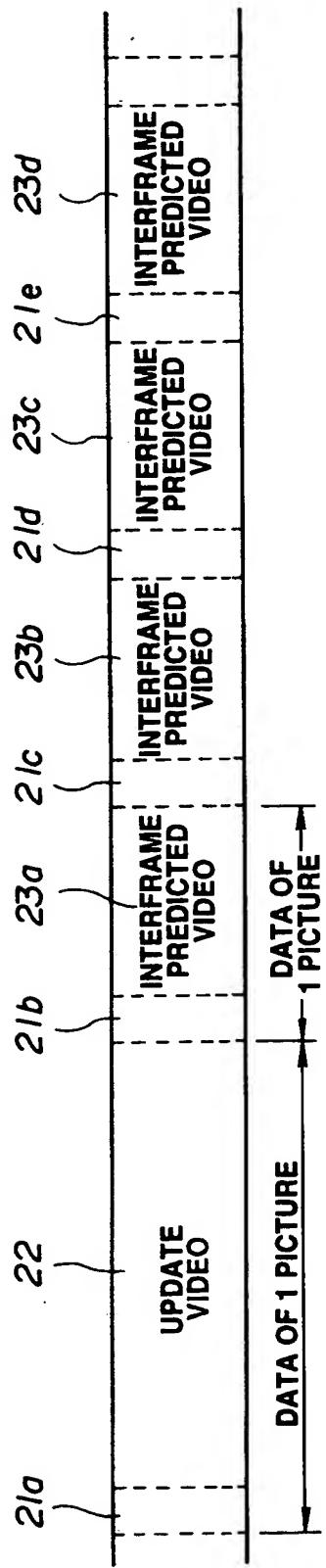
FIG.4
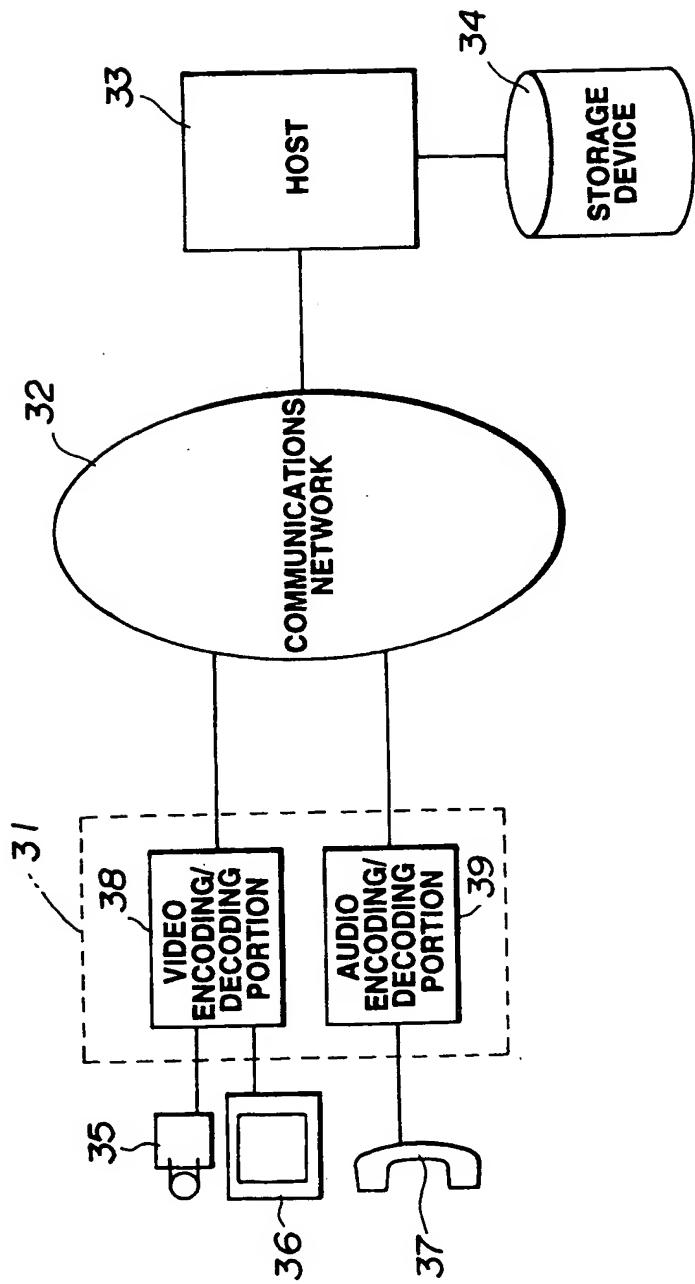
FIG.5
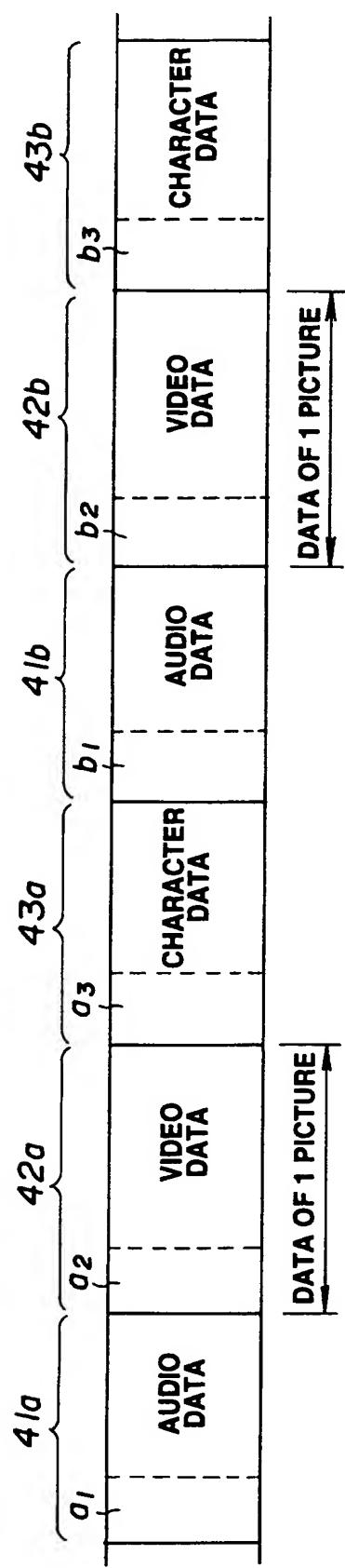
FIG.6
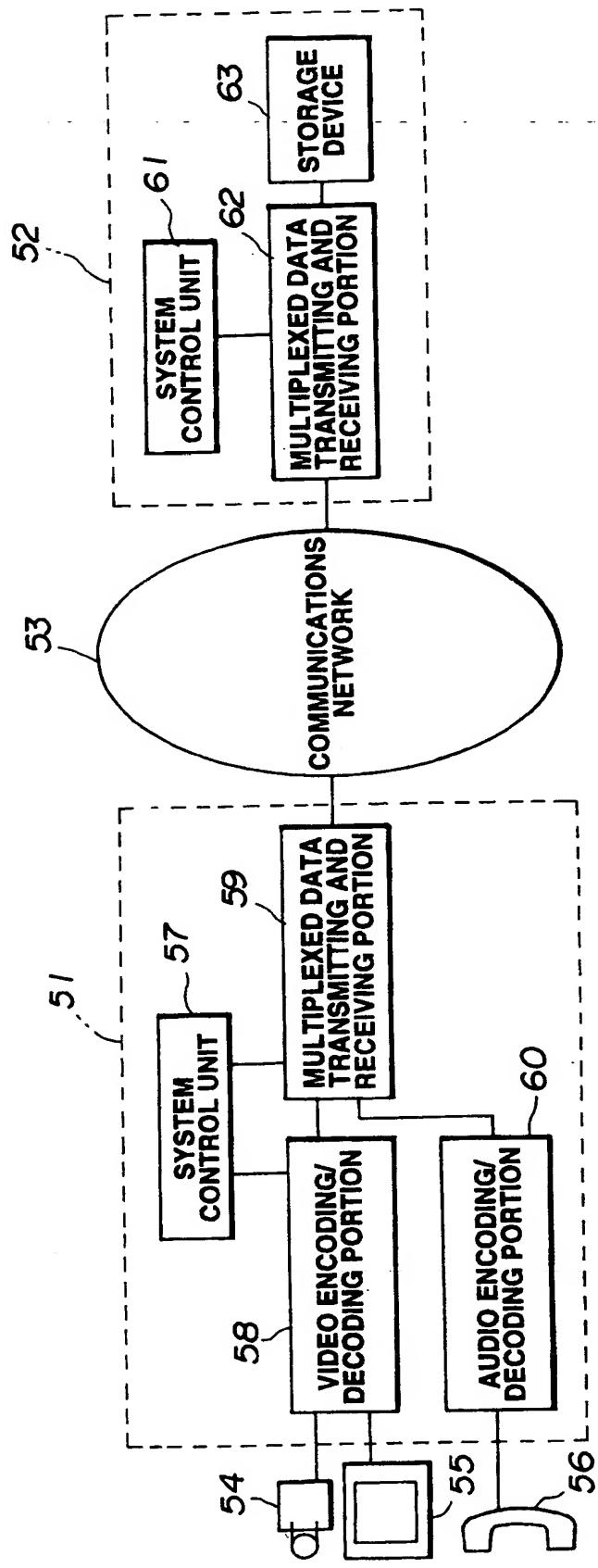
FIG.7
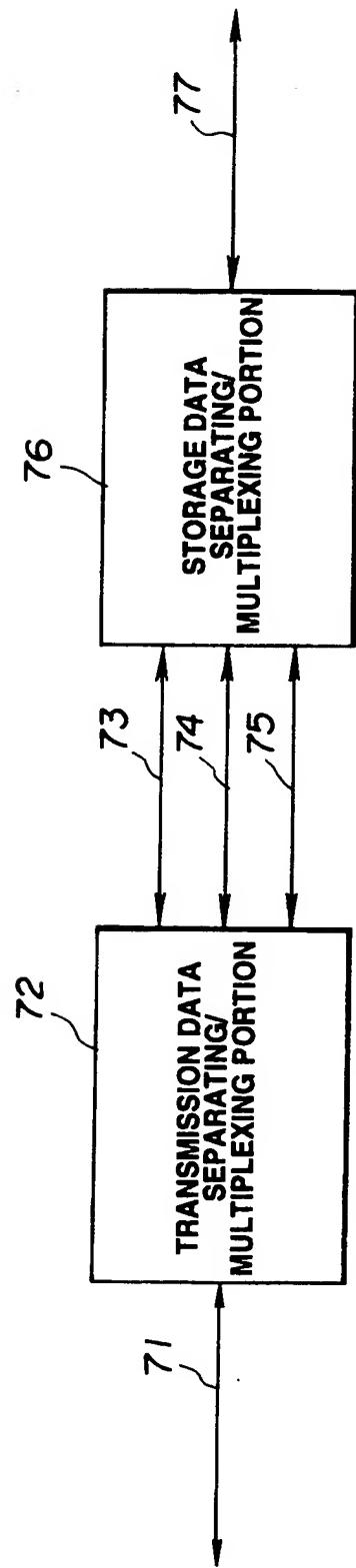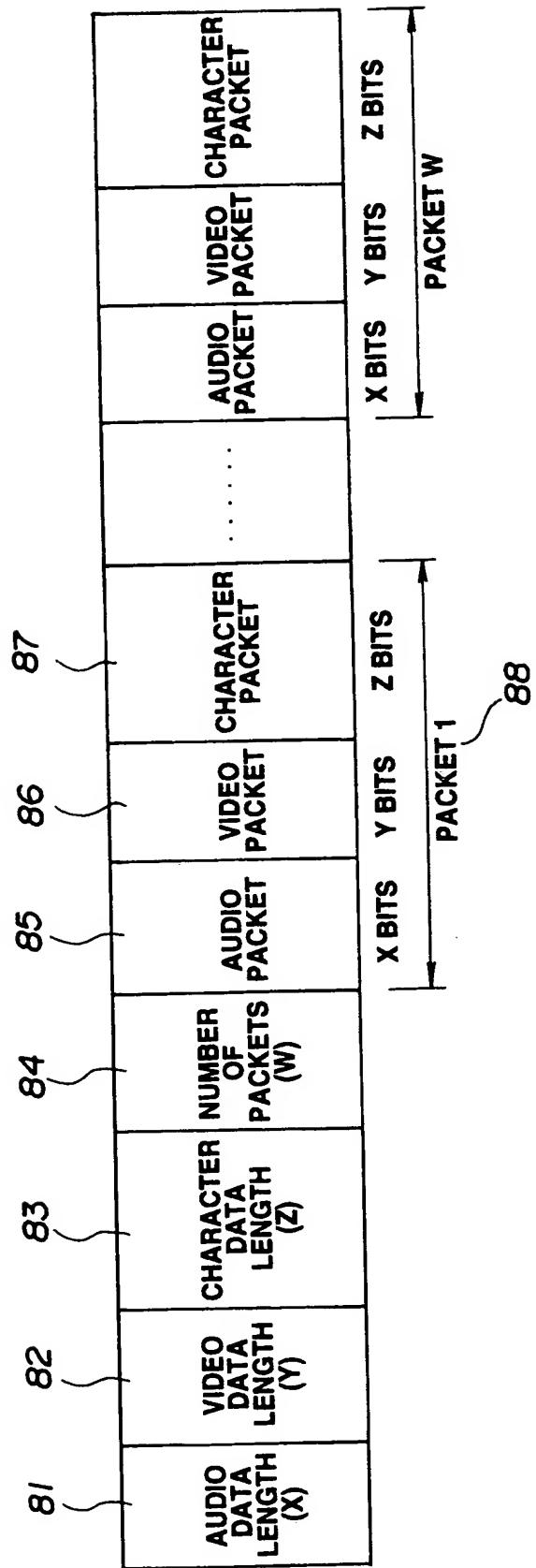
STORAGE DATA SEPARATING/ MULTIPLEXING PORTION

76

73

74

75

TRANSMISSION DATA SEPARATING/ MULTIPLEXING PORTION

72

77

71

FIG.8

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.5)
Y	US-A-5 138 440 (A. M. RADICE) * column 1, line 54 - column 2, line 38 * ---	1-5	H04Q11/04 H04N7/137
Y	REVUE HF TIJDSCHRIFT vol. 15, no. 3/4 , March 1991 , BRUSSELS pages 102 - 112 XP000307855 S. D'AGOSTINO 'Universal Videocodec for ATM Networks' * page 105, line 28 - line 32 * ---	1-5	
A	COMPUTER vol. 18, no. 10 , October 1985 , SILVER SPRING, MARYLAND pages 82 - 89 J. K. REYNOLDS, J. B. POSTEL, A. R. KATZ, G. G. FINN, A. L. DESCHON. 'The DARPA Experimental Multimedia Mail System' -----		
			TECHNICAL FIELDS SEARCHED (Int.Cl.5) H04N H04Q H04M

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	29 March 1994	Veen, G